Sinusoidal coding

The invention relates to encoding a signal, in which frequency and amplitude information of at least one sinusoidal component are determined and sinusoidal parameters representing the frequency and amplitude information are transmitted.

US-A 5,664,051 discloses a speech decoder apparatus for synthesizing a speech signal from a digitized speech bit-stream of the type produced by processing speech with a speech encoder. The apparatus includes an analyzer for processing the digitized speech bit stream to generate an angular frequency and magnitude for each of a plurality of sinusoidal components representing the speech processed by the speech encoder, the analyzer generating the angular frequencies and magnitudes over a sequence of times; a random signal generator for generating a time sequence of random phase components; a phase synthesizer for generating a time sequence of synthesized phases for at least some of the sinusoidal components, the synthesized phases being generated from the angular frequencies and random phase components; and a synthesizer for synthesizing speech from the time sequences of angular frequencies, magnitudes and synthesized phases. This document discloses that a great improvement in the quality of synthesized speech can be achieved by not encoding the phase of harmonics in voiced (i.e., composed primarily of harmonics) portions of the speech, and instead synthesizing an artificial phase for the harmonics at the receiver. By not encoding this harmonic phase information, the bits that would have been consumed in representing the phase are available for improving the quality of the other components of the encoded speech (e.g. pitch, harmonic magnitudes). In synthesizing the artificial phase, the phase and frequencies of the harmonics within the segments are taken into account. In addition, a random phase component, or jitter, is added to introduce randomness in the phase. More jitter is used for speech segments in which a greater fraction of the frequency bands are unvoiced. The random jitter improves the quality of the synthesized speech, avoiding the buzzy, artificial quality that can result when phase is artificially synthesized.

An object of the invention is to provide advantageous coding. To this end, the invention provides a method of encoding a signal, a method of decoding an encoded signal, an audio coder, an audio player, an audio system, an encoded signal and a storage medium as defined in the independent claims. Advantageous embodiments are defined in the dependent claims. The invention provides an advantageous way of applying phase jitter by transmitting a phase jitter parameter from the encoder to the decoder to indicate the amount of phase jitter that should be applied in the decoder during synthesis. Sending a phase jitter parameter has, inter alia, the advantage that a relation between the amount of phase jitter applied in the decoder and the original signal is established. In this way, more natural sound of a reconstructed audio signal is obtained, which better corresponds to the original audio signal. Further, the amount of phase jitter to be applied can be determined faster and more reliable, because it is not necessary to determine locally in the decoder the amount of phase jitter to be applied to generate a natural sounding signal.

By including the phase jitter parameter in the encoded bit-stream, the bit-rate is increased. However, the increase bit-rate can be minimal since these phase jitter parameters can have a very low update-rate, e.g. once per track. A track is a sinusoidal component with a given frequency and amplitude, i.e. a complete set of sinusoid segments. Preferably, the phase jitter parameter is transmitted approximately together with the frequency and the amplitude of the sinusoid at a first instance of a track. In that case, all required information is available at an early stage in the decoding.

An alternative solution to this problem would be to transmit the original phase, or phase differences at various time instances such that the frequency can be adapted during synthesis to match this original phase at the respective time instances. Sending these original phase parameters result in a better quality but requires a higher bit-rate.

In a preferred embodiment, it is assumed that phase-jitter applied to harmonically related frequencies bears the same harmonic relation as the related frequencies. It than suffices to transmit one phase jitter parameter per group of harmonically related frequencies.

The phase jitter parameters are preferably derived from statistical deviations measured in the original phase. In a preferred embodiment, a difference between an original phase of the signal and a predicted phase is determined, which predicted phase is calculated from the transmitted frequency parameters and a phase continuation requirement, and the phase jitter parameter is derived from said difference. With continuous phase, only a first instance of a sinusoid in each track may include a phase parameter, consecutive segments of

the sinusoid must match, i.e. calculate, their phase parameters in such a way that they align with the phase of the current sinusoid segment. Reconstructed phases based on a continuous phase criterion lost their relation to original phases. As explained in the prior art, reconstructed signals with a constant frequency and amplitude in conjunction with continuous phases, sound somewhat artificial.

In general, it is not required that the phase jitter parameters indicate an exact amount of phase jitter. The decoder may perform a certain predetermined calculation based on the value of the phase jitter parameter and/or characteristics of the signal.

In an extreme case, the phase jitter parameter consists of one bit only. In this case, e.g. a zero indicates that no phase jitter should be applied and a one indicates that phase jitter should be applied. The phase jitter to be applied in the decoder may be a predetermined amount or may be derived in a pre-determined manner from characteristics of the signal.

The aforementioned and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter.

In the drawings:

Fig. 1 shows an illustrative embodiment comprising an audio coder according to the invention;

Fig. 2 shows an illustrative embodiment comprising an audio player according to the invention; and

Fig. 3 shows an illustrative embodiment of an audio system according to the invention.

The drawings only show those elements that are necessary to understand the invention.

The invention is preferably applied in a general sinusoidal coding scheme, not only in speech coding schemes, but also in sinusoidal audio coding schemes. In a sinusoidal coding scheme, an audio signal to be encoded is represented by a plurality of sinusoids of which a frequency and an amplitude are determined in an encoder. Often, the phase is not transmitted, but the synthesis is performed in such a way that the phase between two subsequent segments is continuous. This is done to save bit-rate. In a typical sinusoidal coding scheme sinusoidal parameters for a number of sinusoidal components are extracted. The sinusoidal parameter set for one component at least consists of a frequency and an amplitude. More sophisticated coding schemes also extract information on the course of the

frequency and/or amplitude as a function of time. In the simplest case, the frequency and amplitude are assumed to be constant within a certain amount of time. This time is denoted as the update interval and typically ranges from 5ms – 40 ms. During synthesis, the frequencies and amplitudes of consecutive frames have to be connected. A tracking algorithm can be applied to identify frequency tracks. Based on this information, a continuous phase can be calculated such that the sinusoidal components corresponding to a single track properly connect. This is important because it prevents phase discontinuities, which are almost always audible. Since the frequencies are constant over each update interval, the continuously reconstructed phase has lost its relation to the original phase.

Fig. 1 shows an exemplary audio coder 2 according to the invention. An audio signal *A* is obtained from an audio source 1, such as a microphone, a storage medium, a network etc. The audio signal *A* is input to the audio coder 2. A sinusoidal component in the audio signal *A* is parametrically modeled in the audio coder 2. A coding unit 20 derives from the audio signal *A*, a frequency parameter *f* and an amplitude parameter *a* of at least one sinusoidal component. These sinusoidal parameters *f* and *a* are included in an encoded audio signal *A'* in multiplexer 21. The audio stream *A'* is furnished from the audio coder to an audio player over a communication channel 3, which may be a wireless connection, a data bus or a storage medium, etc. At the encoder, a sinusoidal track is identified. This means that at two time instants t_1 and t_2, the frequencies and phase are known. From the frequency track and phase at t_1, the phase at t_2 can be predicted. This is preferably done in a same way as in a decoder. The error of the prediction of the phase at t_2 and the actual measured phase can be calculated. A characteristic value of this error, e.g. mean absolute value or a variance, can be determined. Preferably, the phase jitter parameter is derived from this characteristic value. In this way, the required phase jitter is determined in the encoder, by calculating the difference between the actual phase and the phase determined from the sinusoidal parameters in the encoder. A phase jitter parameter derived from this difference is transmitted to the decoder which uses the phase jitter parameter to introduce a derived amount of phase jitter by changing slightly the phase of the corresponding signal in the synthesis.

An alternative way of determining the phase jitter parameter is to monitor fluctuations in the original frequency.

An embodiment comprising an audio player 4 according to the invention is shown in Fig. 2. An audio signal *A'* is obtained from the communication channel 3 and de-multiplexed in de-multiplexer 40 to obtain the sinusoidal parameters *f* and *a* and the phase jitter parameter *p* that are included in the encoded audio signal *A'*. These parameters *f*, *a* and

p are furnished to a sinusoidal synthesis (SS) unit 41. In SS unit 41, a sinusoidal component S' is generated which has approximately the same properties as the sinusoidal component S in the original audio signal A. The sinusoidal component S' is multiplexed together with other reconstructed components and output to an output unit 5, which may be a loudspeaker. At the decoder, the phase jitter parameter p is available. Next to determining the phase of the signal at each instant by using phase continuation and some way of frequency (and thus phase) interpolation, the phase jitter parameter is used to add a disturbance to the constructed phase interpolation. This new phase is then treated as 'original phase', to the extent that the frequencies are adjusted during synthesis to match these new phase values.

Fig. 3 shows an audio system according to the invention comprising an audio coder 2 as shown in Fig. 1 and an audio player 4 as shown in Fig. 2. Such a system offers playing and recording features. The communication channel 3 may be part of the audio system, but will often be outside the audio system. In case the communication channel 3 is a storage medium, the storage medium may be fixed in the system or may also be a removable disc, tape, memory stick etc.

It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be able to design many alternative embodiments without departing from the scope of the appended claims. In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. The word 'comprising' does not exclude the presence of other elements or steps than those listed in a claim. The invention can be implemented by means of hardware comprising several distinct elements, and by means of a suitably programmed computer. In a device claim enumerating several means, several of these means can be embodied by one and the same item of hardware. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

In summary, encoding a signal is provided, wherein frequency and amplitude information of at least one sinusoidal component in the signal is determined, and sinusoidal parameters representing the frequency and amplitude information are transmitted, and wherein further a phase jitter parameter is transmitted, which represents an amount of phase jitter that should be added during restoring the sinusoidal component from the transmitted sinusoidal parameters.